UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

ESTIMATING THE TOBIT MODEL WITH SERIAL CORRELATION : AN OPERATIONAL APPROACH*

by

Marcel G. Dagenais[1]

[1] Département de sciences économiques, Université de Montréal

* This research project was financed mainly by the Fonds F.C.A.R. du Québec. It was also supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Social Sciences and Humanities Research Council of Canada.

I benefitted from the help of Luc Bussière and Sylvie Hébert as research assistants, in the exploratory phase of the project. Dominique Paskievici contributed significantly to the project as programmer-analyst.

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche.

Université de Montréal

Centre de recherche sur les transports - Publication #452
Département de sciences économiques - Cahier #8615

Décembre 1985

Dépôt légal - 2e trimestre 1986

REPORT DOCUMENTATION FORM

1. Title Estimating the Tobit model with serial correlation : an operational approach.	**2. Report Date** December 1985
	3. C.R.T. Report No. 452
	4. ISBN

5. Author(s)

Marcel G. Dagenais

6. Sponsoring Organization Name and Address

7. Abstract

Several authors have discussed recently the limited dependent variable regression model with serial correlation between residuals. The pseudo-maximum likelihood estimators obtained by ignoring serial correlation altogether, have been shown to be consistent. We present alternative pseudo-maximum likelihood estimators which are obtained by ignoring serial correlation only selectively. Monte Carlo experiments on a model with first order serial correlation suggest that our alternative estimators have substantially lower mean-squared errors in medium size and small samples, especially when the serial correlation coefficient is high. The same experiments also suggest that the true level of the confidence intervals established with our estimators by assuming asymptotic normality, is somewhat lower than the intended level. Although the paper focuses on models with only first order serial correlation, the generalization of the proposed approach to serial correlation of higher order is also discussed briefly.

8. Key Words limited dependent variable; serial correlation; Monte Carlo; regression analysis	**9. No. of Pages : 26**
	10. Price Code : A02

ABSTRACT - RÉSUMÉ

Several authors have discussed recently the limited dependent variable regression model with serial correlation between residuals. The pseudo-maximum likelihood estimators obtained by ignoring serial correlation altogether, have been shown to be consistent. We present alternative pseudo-maximum likelihood estimators which are obtained by ignoring serial correlation only selectively. Monte Carlo experiments on a model with first order serial correlation suggest that our alternative estimators have substantially lower mean-squared errors in medium size and small samples, especially when the serial correlation coefficient is high. The same experiments also suggest that the true level of the confidence intervals established with our estimators by assuming asymptotic normality, is somewhat lower than the intended level. Although the paper focuses on models with only first order serial correlation, the generalization of the proposed approach to serial correlation of higher order is also discussed briefly.

* * * * * *

Plusieurs auteurs ont discuté récemment de l'estimation du modèle de régression avec variable dépendante limitée, lorsqu'il y a autocorrélation des erreurs résiduelles. On a, par exemple, démontré que les estimateurs du pseudo-maximum de vraisemblance obtenus en ignorant totalement l'autocorrélation des erreurs étaient convergents. Nous présentons ici des estimateurs du pseudo-maximum de vraisemblance alternatifs, obtenus en négligeant l'autocorrélation des erreurs seulement de façon sélective. Des expériences de Monté Carlo effectuées sur un modèle où les erreurs suivent un processus autorégressif du premier ordre suggèrent que dans les échantillons de tailles moyenne ou petite, nos estimateurs alternatifs ont des écarts quadratiques moyens nettement plus petits que ceux des estimateurs obtenus en ignorant complètement l'autocorrélation des erreurs. Le gain réalisé s'avère notablement plus grand lorsque le coefficient d'autocorrélation est élevé. Les mêmes expériences suggèrent également que le véritable niveau des intervalles de confiance générés à l'aide de nos estimateurs, en supposant la normalité asymptotique, est un peu inférieur au niveau recherché. Quoique notre exposé traite principalement des modèles qui ne contiennent qu'un processus autorégressif du premier ordre, nous discutons également de façon très brève, de la généralisation possible de notre approche aux modèles dans lesquels les erreurs résiduelles suivent un processus autorégressif d'un ordre plus élevé.

I. INTRODUCTION

Several authors have discussed recently the limited dependent variable model [Tobin (1958)], when there is serial correlation between the residual errors. Robinson (1982) has shown that the Pseudo-Maximum Likelihood estimators obtained by maximizing the log-likelihood function which Ignores the Serial correlation (PMLIS) are consistent. Robinson shows also that these estimators are asymptotically normally distributed and he derives their asymptotic covariance matrix. Dagenais (1982) has derived an expression for the likelihood function of the Tobit model with first order serial correlation, which illustrates that the computation of this function involves the evaluation of multivariate normal integrals. The dimensions of these integrals correspond to the length of the runs of limit observations. Tests for serial dependence in such models have also been discussed by Gourieroux *et al.* (1982) and Robinson *et al.* (1985). Robinson *et al.* (1985) note also that the above mentioned PMLIS estimators obtained for Tobit models with serial correlation have relatively large mean-squared errors, in finite samples. Since, on the other hand, the computation of true maximum likelihood estimators is not an operational procedure in that case, unless the sample contains only short runs of limit observations, it would be desirable to have on hand an alternative estimating procedure which would be computationally no more difficult than the PMLIS approach mentioned above, but could reduce substantially the variability of the resulting estimators.

The purpose of this paper is to suggest such a procedure for Tobit models with first order serial correlation. Starting from the true likelihood function of the model and viewing the likelihood which ignores serial correlation as an approximation to the true function, a different approximation is then derived and Alternative Pseudo-Maximum Likelihood (APML) estimators are proposed, in Section II, which maximize this latter function. Since the asymptotic properties of the APML estimators appear to be difficult to establish, Monte Carlo experiments are performed to compare these alternative estimators to the PMLIS estimators, in Section III. Using the estimators of the classical multiple regression model with serial correlation [Beach and MacKinnon (1978)] as a yardstick, it is found that these new estimators have substantially smaller mean-squared errors than the PMLIS estimators. Finally, in Section IV, an estimator of the asymptotic covariance matrix of the APML estimators is proposed. The Monte Carlo experiments suggest that the true confidence level of the intervals established with the above covariance matrix, assuming asymptotic normality of the APML estimators, are somewhat lower than the selected level, in medium size samples. In conclusion, the extension of the proposed procedure to models involving second or higher order serial correlation is very briefly discussed.

II. THE SUGGESTED PROCEDURE

Let us assume the following model :

$$(1) \qquad Y_t = X_t\beta + u_t \ , \qquad \text{if } X_t\beta + u_t = I_t > L_t \ ,$$

$$(2) \qquad Y_t = L_t \ , \qquad \text{if } X_t\beta + u_t = I_t \leq L_t \ ,$$

$$(3) \qquad u_t = \rho\, u_{t-1} + \varepsilon_t \qquad (-1 < \rho < 1;\ t = 1,\ldots,N) \ ,$$

where Y_t is the limited dependent variable, X_t is a $1\times K$ vector of exogenous explanatory variables, I_t is a latent variable, L_t is an observed exogenous variable and β is a $K\times 1$ vector of parameters. Furthermore, the ε's are $N(0, \sigma_\varepsilon^2)$ random variables and the u's are normally distributed with zero mean and covariance matrix :

$$(4) \qquad E(uu') = \sum = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho^{N-1} \\ \rho & 1 & & \rho^{N-2} \\ \vdots & \vdots & & \vdots \\ \rho^{N-1} & \rho^{N-2} & \cdots & 1 \end{pmatrix} = \sigma^2\, \Omega \ ,$$

where $u' = (u_1,\ldots,u_N)$ and $\sigma^2 = \sigma_\varepsilon^2/(1-\rho^2)$.

The unknown parameters to be estimated are the elements of the column vector θ, where $\theta' = (\beta', \sigma^2, \rho)$.

As mentioned above, our suggested approach consists, as a first step, in obtaining for the true likelihood function of the model a "closer" approximation than the likelihood under serial independence. The procedure will be easier to expound if we use a simple example. Let us assume that we

have a sample of 5 observations, with the following configuration :

observation numbers	:	1	2	3	4	5
type of observations	:	L	N	L	L	N

where L designates a limit observation and N designates a nonlimit observation. The log-likelihood function of the model may be expressed as :

$$(5) \qquad \tilde{\mathcal{L}} = f(Y_1, Y_2, Y_3, Y_4, Y_5)$$

where $f(Y_1, Y_2, Y_3, Y_4, Y_5)$ is the joint density of $Y_1,\ldots,Y_5$. This density can also be written as follows :

$$(6) \qquad f(Y_1,\ldots,Y_5) = f(Y_2)\ f(Y_1|Y_2)\ \ f(Y_5|Y_1,Y_2)\ f(Y_3,Y_4|Y_1,Y_2,Y_5) \quad .$$

Since the u's obey a first-order Markov process, it can be shown that :

$$(7) \qquad f(Y_5|Y_1,Y_2) = f(Y_5|Y_2) \quad ,$$

$$f(Y_3,Y_4|Y_1,Y_2,Y_5) = f(Y_3,Y_4|Y_2,Y_5) \quad .$$

Hence, the likelihood can be expressed as :

$$(8) \qquad \tilde{\mathcal{L}} = f(Y_2)\ f(Y_1|Y_2)\ f(Y_5|Y_2)\ f(Y_3,Y_4|Y_2,Y_5) \quad .$$

Since Y_2 and Y_5 are nonlimit values, $f(Y_2)$ is a marginal univariate normal density and $f(Y_5|Y_2)$ is a conditional univariate normal density. In addition, $f(Y_1|Y_2)$ is the integral of a conditional univariate normal density; and $f(Y_3,Y_4|Y_2,Y_5)$ is a double integral of a conditional bivariate normal density, since :

$$(9) \qquad f(Y_3,Y_4|Y_2,Y_5) = \int_{-\infty}^{L_3} \int_{-\infty}^{L_4} f(I_3,I_4|Y_2,Y_5)\ dI_4\ dI_3 \quad ,$$

where $f(I_3, I_4 | Y_2, Y_5)$ is the conditional density of (I_3, I_4) given (I_2, I_5) or equivalently given (Y_2, Y_5), since $I_2 = Y_2$ and $I_5 = Y_5$.

One possible simplification of $\tilde{\mathcal{L}}$ is to replace it by :

$$\tilde{\mathcal{L}}^* = f(Y_2)\, f(Y_1)\, f(Y_5)\, f(Y_3)\, f(Y_4) \quad , \tag{10}$$

where $f(Y_2)$ and $f(Y_3)$ are marginal univariate normal densities and $f(Y_1)$, $f(Y_3)$ and $f(Y_4)$ are integrals of marginal univariate normal densities.

The PMLIS estimators of (β, σ^2) are obtained by maximizing $\mathcal{L}^* = \ln \tilde{\mathcal{L}}^*$. A "closer" approximation to $\tilde{\mathcal{L}}$ can be obtained as follows :

$$\tilde{\mathcal{L}}^{**} = f(Y_2)\, f(Y_1|Y_2)\, f(Y_5|Y_2)\, f(Y_3|Y_2,Y_5)\, f(Y_4|Y_2,Y_5) = \prod_{i=1}^{5} \tilde{\mathcal{L}}_i^{**} \quad , \tag{11}$$

where the first three terms on the R.H.S. of $\tilde{\mathcal{L}}^{**}$ are identical with those of $\tilde{\mathcal{L}}$. The last term on the R.H.S. of $\tilde{\mathcal{L}}$, however, is replaced in $\tilde{\mathcal{L}}^{**}$, by the products of two one-dimensional integrals, as if the covariance between Y_3 and Y_4, in the conditional density of (Y_3, Y_4) given (Y_2, Y_5), were equal to zero. Our APML estimators are obtained by maximizing $\mathcal{L}^{**} = \ln \tilde{\mathcal{L}}^{**}$, with respect to θ.

The procedure outlined above for deriving $\tilde{\mathcal{L}}^{**}$ in the five observation example is easy to generalize for any sample size. The terms appearing in $\tilde{\mathcal{L}}^{**}$ are :

a) the marginal density of the first nonlimit observation;

b) the conditional density of every other nonlimit observations, given the preceding nonlimit observation;

c) for every limit observation *preceding* the first nonlimit observation, the integral of the conditional univariate density of this limit observation given the first nonlimit observation;

d) for every other limit observation, the integral of the conditional univariate density of this limit observation, given the preceding nonlimit observation and also the following one (if there is such an observation).

The computation of $\mathcal{L}^{**}$ involves only one-dimensional normal integrals. The evaluation of $\mathcal{L}^{**}$ is therefore hardly more costly than that of $\mathcal{L}^{*}$.

III. MONTE CARLO EXPERIMENTS

Given that, on the one hand, it would be difficult to analyze the asymptotic properties of our APML estimators and that, on the other hand, finite sample properties are of utmost interest, Monte Carlo experiments were performed to examine the biases and the mean-squared errors of our estimators. Comparisons are made with the corresponding properties of the PMLIS estimators, as well as with those of the estimators of the classical multiple regression model with serial correlation, which corresponds to the *unrealistic* situation where the latent variable I_t would be known for all observations. Clearly, since these last estimators [designated henceforth as Maximum Likelihood estimators with Latent Variables (MLLV)] are based on a greater amount of information, they are expected to perform better than our APML estimators.

It is found in all cases that, using the MSE as a criterion, the APML estimators perform better than the PMLIS estimators and that, in fact, their performance is surprisingly good in comparison to that of the MLLV estimators.

A. The Data

In order to simulate conditions which resemble those encountered in economic analyses, the model and the data used were derived from an empirical study performed by Dagenais (1964).

The model is a flexible accelerator model [Lovell (1961)] :

$$(12) \qquad I_t = .129 + .92\, C_t + .26\, C_{t-1} - .77\, M_{t-1} - .75\, B_{t-1} + u_t \ ,$$

$$S_t = I_t \ , \qquad \text{if } I_t < K_t \ ,$$

$$S_t = K_t \ , \qquad \text{if } I_t \geq K_t \ ,$$

$$u_t = \rho\, u_{t-1} + \varepsilon_t \ ,$$

where :

S_t = shipments of North American newsprint mills, during half-year t, in million short tons;

C_t = newsprint consumption of the customers of the North American newsprint mills, in half year t, in million tons;

M_{t-1} = end of period stocks, at the mills, in million tons;

B_{t-1} = end of period stocks, at the customers', in million tons;

K_t = capacity of North American newsprint mills, in period t.

Note that in this application the limit considered is an upper limit. The data can be found in the Appendix[1]. The sample covers the period 1922, first semester to 1960, second semester and contains 76 observations. The values of the β parameters shown in equation (12) are those obtained by Bussière (1983), from the same data, with a slightly different model. The values obtained by Bussière for ρ and σ^2 were .7 and 5700, respectively. In the actual sample, there are 20 limit observations corresponding to the following semiannual periods : 1943-I to 1945-II[2], 1947-I to 1949-I, 1950-II to 1952-I and 1954-II to 1956-II.

B. Design of the Experiments

The main set of experiments was made with the 76 observations shown in the Appendix. For each experiment, σ^2 was set at a given value and so was ρ. Then the series of serially correlated u's were drawn at random and the values of the latent variables were computed. Then, the values of S_t were set at min (I_t, K_t).

---

[1]. The data is similar to the actual data found in Dagenais (1964) except for two observations, namely those referring to 1937, 2nd semester and to 1938, 1st semester. For reasons explained in Dagenais (1964), these observations were very abnormal and could be considered as outliers. They were replaced by more normal values, for performing the present experiments.

[2]. During World War II, the upper limit imposed on production did not actually correspond to the physical capacity of the mills but to quotas imposed by the Canadian and U.S. governments.

For each experiment, a method combining the control variate and the two-antithetic-variate techniques was adopted, as suggested by Mikhail (1972). One hundred samples were first drawn. Then for each sample, a corresponding sample with the u vector replaced by -u was generated. For every sample, the PMLIS, APML and MLLV estimators were computed. Control variates were obtained for the estimators of β and σ^2 by using MLLV estimators, assuming ρ to be known. A control variate was also obtained for the estimator of ρ by assuming the u's to be known and applying ordinary least squares to equation (3).

Six different experiments (designated as experiments E-1 to E-6) were made for ρ = .5, .7, .95 and R^2 = .5, .95, where R^2 corresponds to the theoretical coefficient of determination of the following regression model :

$$Y^* = X^*\beta + \varepsilon \quad , \tag{13}$$

where $Y^{*\prime} = (Y_1\sqrt{1-\rho^2},\ Y_2 - \rho Y_1, \ldots, Y_N - \rho Y_{N-1})$

$X^{*\prime} = (X_1'\sqrt{1-\rho^2},\ X_2' - \rho X_1', \ldots, X_N' - \rho X_{N-1}')$.

R^2 is defined as :

$$R^2 = \beta' x^{*\prime} x^* \beta / (\beta' x^{*\prime} x^* \beta + N\sigma_\varepsilon^2) \quad ,$$

where x^* corresponds to X^* with the variables expressed in mean deviation form.

Table 1 gives the biases of the different parameters for the three estimation procedures considered, namely PMLIS, APML and MLLV, when the sample size is 76, for different values of ρ and R^2.

In order to perform our experiments, the model was reparametrized in terms of $w = \ln \sigma^2$ and $z = \frac{1}{2} \ln [(1+\rho)/1-\rho)]$, to avoid the recourse to constrained maximization procedures. Our results concerning σ^2 and ρ are therefore given in terms of w and z, in Table 1.

One notices that the biases of the β parameters are virtually null in the case of the MLLV procedure except for β_o, when $\rho = .95$. As $\rho \to 1$, β_o becomes indeterminate in the multiple regression model with serial correlation. Hence, it is not surprising that the results become less accurate, when ρ is high. Furthermore, the estimators of z and w are biased downwards and in some cases, this bias is not negligible.

The biases associated with the PMLIS procedure are often relatively large. We have underlined once all cases for which the bias is larger, in absolute value, than 10% of the corresponding parameter. Cases for which this percentage was over 25% have been underlined twice and cases for which the percentage was over 50% were boxed. The PMLIS biases are in many cases much larger than the corresponding APML biases, as can be verified readily from the |PMLIS|/|APML| column, which gives the absolute value of the ratios of the two biases. Only for the B_{t-1} variable are the PMLIS biases smaller than the APML biases, in three cases. These cases are identified by a * in Table 1. But the APML biases are themselves relatively negligible, in two of these cases, being smaller than 1% of the value of the parameter.

TABLE I

BIASES (EXPERIMENTS E-1 TO E-6)

E-1			$\rho = .5$ $R^2 = .5$		
Variables	Parameters values	PMLIS	APML	$\frac{\lvert PMLIS \rvert}{\lvert APML \rvert}$	MLLV
constant	β_o=129.	-459.3836	49.6554	9.25	-0.7870
C_t	β_1=.92	-0.4913	0.0499	9.85	0.0013
C_{t-1}	β_2=.26	0.4493	-0.0296	15.18	-0.0011
M_{t-1}	β_3=-.77	1.8532	-0.2174	8.52	-0.0001
B_{t-1}	β_4=-.75	0.0874	-0.0533	1.64	-0.0004
(1)	z =.55	N.A.	-0.2447	N.A.	-0.0890
	w =12.47	-0.6145	-0.2081	2.95	-0.1511

E-2			$\rho = .5$ $R^2 = .9853$		
Variables	Parameters values	PMLIS	APML	$\frac{\lvert PMLIS \rvert}{\lvert APML \rvert}$	MLLV
constant	β_o=129.	-19.7562	1.5172	13.02	-0.0000
C_t	β_1=.92	-0.0281	0.0023	12.22	0.0000
C_{t-1}	β_2=.26	-0.0251	-0.0023	10.91	0.0000
M_{t-1}	β_3=-.77	0.1036	-0.0114	9.09	0.0000
B_{t-1}	β_4=-.75	0.0017	0.0027	0.63*	-0.0000
(1)	z =.55	N.A.	-0.1628	N.A.	-0.1143
	w =8.27	-0.2866	-0.2279	1.26	-0.1773

E-3			$\rho = .7$ $R^2 = .5$		
constant	β_o=129.	-343.6106	73.5197	4.67	2.8688
C_t	β_1=.92	-0.3453	0.0938	3.68	0.0007
C_{t-1}	β_2=.26	0.3237	-0.0718	4.51	-0.0010
M_{t-1}	β_3=-.77	1.4469	-0.2988	4.84	-0.0062
B_{t-1}	β_4=-.75	0.0332	-0.0511	0.65*	-0.0015
(1)	z =.8673	N.A.	-0.2875	N.A.	-0.1488
	w =11.99	-0.6609	-0.3630	1.82	-0.2684

E-4			$\rho = .7$ $R^2 = .9658$		
constant	β_o=129.	-24.3703	1.1084	21.99	-0.0000
C_t	β_1=.92	-0.0366	0.0042	8.71	1.0000
C_{t-1}	β_2=.26	0.0322	-0.0019	17.47	0.0000
M_{t-1}	β_3=-.77	0.1309	-0.0163	8.03	0.0000
B_{t-1}	β_4=-.75	-0.0004	-0.0016	0.25*	0.0000
(1)	z =.8673	N.A.	-0.2082	N.A.	-0.1544
	w =8.65	-0.4426	-0.3507	1.26	-0.2762

E-5			$\rho = .95$ $R^2 = .5$		
constant	β_o=129.	-272.172	103.527	2.63	2.657
C_t	β_1=.92	-0.249	0.026	9.58	0.000
C_{t-1}	β_2=.26	0.227	-0.025	9.08	0.000
M_{t-1}	β_3=-.77	1.048	-0.095	11.03	0.002
B_{t-1}	β_4=-.75	0.089	-0.063	1.41	0.000
(1)	z =1.8318	N.A.	-0.454	N.A.	-0.387
	w =12.15	-1.502	-0.865	1.74	-0.770

E-6			$\rho = .95$ $R^2 = .98$		
constant	β_o=129.	-8.5195	0.8949		-0.1913
C_t	β_1=.92	-0.0146	-0.0003		-0.0000
C_{t-1}	β_2=.26	0.0123	0.0003		-0.0000
M_{t-1}	β_3=-.77	0.0518	0.0007		-0.0001
B_{t-1}	β_4=-.75	0.0018	0.0002		0.0001
(1)	z =1.8318	N.A.	-0.4315	N.A.	-0.4256
	w =8.26	-1.2709	-0.8834	1.44	-0.8544

(1) For computational purposes, the model was reparametrized in terms of $w = \ln c^2$ and $z = .5 \ln [(1+\rho)/(1-\rho)]$.

$|PMLIS|/|APML|$ corresponds to the absolute value of the ratio of the PMLIS bias divided by the APML bias.

* indicates the cases where the APML bias is larger than the PMLIS bias, in absolute value.

_____ : bias larger than 10% of the parameter value.

_____ : bias larger than 25% of the parameter value.

▭ : bias larger than 50% of the parameter value.

The APML approach underestimates w and z in all cases and the biases are large in some cases, for z. Note, however, that in four out of six cases, these biases are not very much larger than those found for MLLV, which does yield consistent estimators for w and z.

Table II gives, for the same experiments, the ratios of the mean-squared errors of the three estimators considered, for each parameter. The APML estimators outperform the PMLIS estimators in all cases. The relative advantage of the APML procedure over the PMLIS procedure increases as the serial correlation gets higher. The PMLIS approach produces especially relatively large mean-squared errors when $\rho = .95$. In contrast, the APML estimators do not yield worse results, in comparison to the MLLV estimators, as ρ increases from .5 to .95. Given that the MLLV estimators use information that is actually *unavailable* in situations where the Tobit model is applied, —since the MLLV estimators require the knowledge of the latent variables which, in reality, are *unobservable* for the limit observations,— the fact that the ratio of the MSE of the MLLV estimators over the MSE of the MLLV estimators is, on average, of the order of 1.75, even when $\rho = .95$, and $R^2 = .5$, and of the order of 1.25 when $\rho = .95$ and $R^2 = .98$, appears as a very good performance.

RATIOS OF MEAN-SQUARED ERRORS FOR EXPERIMENTS E-1 - E-6

	E-1 $\rho = .5$ $R^2 = .5$			E-2 $\rho = .5$ $R^2 = .9853$		
	$\frac{MSE(PMLIS)}{MSE(MLLV)}$	$\frac{MSE(APML)}{MSE(MLLV)}$	$\frac{MSE(PMLIS)}{MSE(APML)}$	$\frac{MSE(PMLIS)}{MSE(MLLV)}$	$\frac{MSE(APML)}{MSE(MLLV)}$	$\frac{MSE(PMLIS)}{MSE(APML)}$
constant: β_o	2.48	1.21	2.05	1.44	1.14	1.27
C_t : β_1	2.90	1.29	2.25	2.00	1.11	1.80
C_{t-1} : β_2	2.61	1.28	2.05	2.33	1.19	1.96
M_{t-1} : β_3	4.55	1.46	3.12	2.31	1.23	1.87
B_{t-1} : β_4	1.34	1.32	1.01	1.79	1.32	1.36
z	N.A.	3.10	N.A.	N.A.	1.67	N.A.
w	6.37	1.43	4.44	1.77	1.41	1.25
	E-3 $\rho = .7$ $R^2 = .5$			E-4 $\rho = .7$ $R^2 = .9658$		
constant: β_o	1.94	1.42	1.37	1.46	1.08	1.34
C_t : β_1	3.97	1.77	2.24	3.12	1.15	2.71
C_{t-1} : β_2	4.08	1.83	2.23	4.18	1.17	3.58
M_{t-1} : β_3	6.81	1.64	4.15	4.06	1.28	3.18
B_{t-1} : β_4	1.96	1.48	1.33	3.10	1.27	2.45
z	N.A.	2.67	N.A.	N.A.	1.57	N.A.
w	3.60	1.50	2.40	1.83	1.38	1.33
	E-5 $\rho = .95$ $R^2 = .5$			E-6 $\rho = .95$ $R^2 = .98$		
constant: β_o	1.23	1.33	0.92	1.26	1.03	1.23
C_t : β_1	11.24	1.78	6.30	9.41	1.29	7.30
C_{t-1} : β_2	10.68	1.99	5.36	10.62	1.35	7.87
M_{t-1} : β_3	38.10	2.17	17.55	39.45	1.17	33.73
B_{t-1} : β_4	12.18	2.04	5.96	15.26	1.38	11.05
z	N.A.	1.49	N.A.	N.A.	1.06	N.A.
w	2.22	1.35	1.64	11.90	1.05	11.34

Table III gives, for each of the six experiments concerned, the percentage of limit and nonlimit observations contained in the two hundred samples generated. Table III also gives, for the limit observations, the cumulative distribution of observations contained in runs of different length.

A simple inspection of Table III indicates that the average length of the runs of limit observations tends to increase with ρ and decrease with R^2. In experiment E-5 which is characterized by $\rho = .95$ and $R^2 = .5$, the samples that were generated contained runs of over 30 limit observations. The fact that our procedure still performed well in such an experiment is reassuring. The PMLIS estimator performed rather badly in this case, as can be seen from Table II.

The data used for the explanatory variables in the experiments reported in Tables I and II were highly collinear, since the determinant of their correlation matrix is equal to 0.004. We therefore repeated one of the above experiments, namely E-4, after having orthogonalized the data matrix and modified the β parameters accordingly. The results of this new experiment (E-7) are reported in Table IV. In addition, Table IV gives the results of two further experiments. Experiment E-8 also repeats E-4, with the sample reduced to 25 observations and experiment E-9 corresponds again to E-4, but with ρ set to -.7 instead of +.7. In E-7, the percentage of limit observations was 24%, and 17% of the observations were limit observations contained in runs of length greater than 4. In experiments E-8 and E-9, the corresponding percentages were 34% and 15% for E-8, 24% and 8% for E-9.

TABLE III

PERCENTAGE OF OBSERVATIONS CONTAINED IN RUNS OF DIFFERENT LENGTHS, FOR EXPERIMENTS REPORTED IN TABLES I AND II

	E-1 $\rho = .5$ $R^2 = .5$	E-2 $\rho = .5$ $R^2 = .9853$	E-3 $\rho = .7$ $R^2 = .5$	E-4 $\rho = .7$ $R^2 = .9658$	E-5 $\rho = .95$ $R^2 = .5$	E-6 $\rho = .95$ $R^2 = .98$
percentage of nonlimit observations	70.20	76.20	71.73	75.77	71.01	76.35
percentage of limit observations contained in runs of length greater than or equal to : 1	29.80	23.80	28.27	24.03	28.99	23.65
2	24.56	21.79	24.92	22.24	27.68	21.96
3	19.19	19.22	21.64	19.81	26.28	19.95
4	14.57	16.67	19.13	16.91	25.00	16.85
5	11.02	11.88	16.05	13.33	23.63	13.88
6	8.26	6.75	14.01	9.38	22.91	10.06
7	6.68	3.39	11.29	6.42	21.41	6.15
8	5.02	2.56	9.13	5.06	20.17	5.41
9	3.60	2.14	7.50	4.55	19.12	5.25
10	2.71	1.61	7.37	3.96	17.82	5.19
11	1.79	1.16	4.79	3.10	16.57	4.99
12	1.50	0.57	3.63	2.01	15.12	3.61
13	1.11	0.10	3.16	1.06	13.62	2.43
14	0.93	0.10	2.82	0.97	12.94	2.43
15	0.75	0.10	2.64	0.69	12.48	2.43
16	0.45	0.00	2.34	0.39	11.59	2.03
17	0.13	0.00	1.81	0.39	10.22	2.03
18	0.13	0.00	1.70	0.39	9.44	1.92
19	0.13	0.00	1.46	0.27	8.73	1.33
20	0.00	0.00	1.08	0.14	6.98	1.20
21	0.00	0.00	0.95	0.14	6.45	1.20
22	0.00	0.00	0.67	0.00	5.48	0.51
23	0.00	0.00	0.67	0.00	5.33	0.36
24	0.00	0.00	0.52	0.00	5.33	0.36
25	0.00	0.00	0.36	0.00	5.01	0.36
26	0.00	0.00	0.36	0.00	5.01	0.36
27	0.00	0.00	0.19	0.00	4.67	0.36
28	0.00	0.00	0.19	0.00	4.31	0.18
29	0.00	0.00	0.19	0.00	3.39	0.00
30	0.00	0.00	0.00	0.00	1.67	0.00
31	0.00	0.00	0.00	0.00	1.27	0.00
32	0.00	0.00	0.00	0.00	0.66	0.00
33	0.00	0.00	0.00	0.00	0.66	0.00
34	0.00	0.00	0.00	0.00	0.23	0.00
35	0.00	0.00	0.00	0.00	0.23	0.00
36	0.00	0.00	0.00	0.00	0.00	0.00

TABLE IV

RATIOS OF MEAN-SQUARED ERRORS FOR EXPERIMENTS E-7 TO E-9

	E-7 : orthogonal matrix			E-8 : N = 25			E-9 : $\rho = -.7$		
	$\frac{MSE(PMLIS)}{MSE(MLLV)}$	$\frac{MSE(APML)}{MSE(MLLV)}$	$\frac{MSE(PMLIS)}{MSE(APML}$	$\frac{MSE(PMLIS)}{MSE(MLLV)}$	$\frac{MSE(APML)}{MSE(MLLV)}$	$\frac{MSE(PMLIS)}{MSE(APML)}$	$\frac{MSE(PMLIS)}{MSE(MLLV)}$	$\frac{MSE(APML)}{MSE(MLLV)}$	$\frac{MSE(PMLIS)}{MSE(APML)}$
constant: β_o	1.56	1.21	1.29	1.51	1.39	1.09	7.07	1.08	6.54
C_t : β_1	5.07	1.38	3.67	2.07	1.40	1.48	4.55	1.30	3.50
C_{t-1} : β_2	4.88	1.21	4.01	1.71	1.31	1.30	4.64	1.41	3.29
M_{t-1} : β_3	2.71	1.41	1.92	1.75	1.30	1.34	11.43	1.28	8.90
B_{t-1} : β_4	1.30	1.16	1.12	1.47	1.47	1.00	3.68	1.62	2.27
z	N.A.	1.60	N.A.	N.A.	1.58	1.29	N.A.	1.02	N.A.
w	1.78	1.32	1.35	2.01	1.38	1.45	1.31	1.03	1.27

The results of these three experiments are very similar to those of the first 6 experiments.

The APML estimators perform rather well in comparison with the MLLV estimators and they outperform the PMLIS estimators systematically. In particular, the MSE of the PMLIS estimators of the β's are relatively large for experiment E-9 with the negative ρ.

IV. SETTING CONFIDENCE ITERVALS

An estimate of the asymptotic covariance matrix of the APML estimators can also be obtained. Following the procedure suggested by Goldberger, Naga and Odeh (1961), this asymptotic covariance estimator may be expressed as :

$$\hat{V}(\hat{\theta}) = \left(\frac{\partial^2 \mathcal{L}^{**}}{\partial\theta\partial\theta'}\right)^{-1}_{\theta=\hat{\theta}} \hat{V}\left(\frac{\partial \mathcal{L}^{**}}{\partial\theta}\right) \left(\frac{\partial^2 \mathcal{L}^{**}}{\partial\theta\partial\theta'}\right)^{-1}_{\theta=\hat{\theta}} , \tag{14}$$

where $\hat{\theta}$ is the APML estimator of θ, $\hat{V}(\hat{\theta})$ is its asymptotic covariance matrix estimator and $\hat{V}\left(\frac{\partial \mathcal{L}^{**}}{\partial\theta}\right)$ is an estimator of the asymptotic covariance matrix of the first derivatives of $\mathcal{L}^{**}$ with respect to the elements of θ. In turn, $\hat{V}\left(\frac{\partial \mathcal{L}^{**}}{\partial\theta}\right)$ can be evaluated as follows :

$$\hat{V}\left(\frac{\partial \mathcal{L}^{**}}{\partial\theta}\right) = \left(\sum_{j=1}^{G} A_j\right)_{\theta=\hat{\theta}} ,$$

where $A_j = \left(\sum_{i\epsilon g_j} \frac{\partial \mathcal{L}_i^{**}}{\partial\theta}\right)\left(\sum_{i\epsilon g_j} \frac{\partial \mathcal{L}_i^{**}}{\partial\theta}\right)'$,

where g_j corresponds to the j'th group of observations. These groups are defined with reference to the $\tilde{\mathcal{L}}$ function shown in equation (8). All observations for which the conditional probabilities are not separable, in $\tilde{\mathcal{L}}$, are elements of the same group. The total number of groups is equal to G.

Using $\hat{V}(\hat{\theta})$, 95% confidence intervals were obtained for each parameter, for each of the samples generated in the first 6 experiments described above. Then for each experiment, the proportions of confidence intervals that contained the true parameters were calculated and are reported in Table V. As can be verified from this Table, the confidence intervals obtained for both the APML and the MLLV estimators are somewhat biased downwards and somewhat erratic, with the APML estimators performing again almost as well as the MLLV estimators. The cases where the downward biases of the confidence intervals are the most important are those relating to the constant term in experiments E-5 and E-6.

TABLE V

ACTUAL LEVELS OF COMPUTED 95% CONFIDENCE INTERVALS

variables	E-1		E-2		E-3		E-4		E-5		E-6	
	APML	MLLV	APML	MLLV	APML	MLLV	APML	MLLV	APML	MLLV	APML	MLLV
constant: β_o	92.5	90.5	87.0	92.0	87.5	88.0	85.0	86.0	67.0	67.5	76.5	78.0
C_t : β_1	92.5	94.0	97.0	95.0	91.0	94.0	97.5	96.0	85.5	91.0	88.5	86.0
C_{t-1} : β_2	95.0	94.0	95.0	94.0	94.0	96.0	95.5	94.0	86.5	88.0	92.5	94.0
M_{t-1} : β_3	91.0	94.0	92.0	91.0	95.5	92.0	90.5	91.0	92.5	96.0	88.0	94.0
B_{t-1} : β_4	90.5	90.5	87.0	89.0	85.0	88.0	91.5	95.0	95.0	94.0	92.0	95.0
average level	92.3	92.6	91.6	92.2	90.6	91.6	92.0	92.4	85.3	87.3	87.5	89.4
averave level excluding the constant	92.3	93.1	92.8	92.3	91.4	92.5	93.8	94.0	89.9	92.3	90.3	92.3
MSE *	9.8	8.7	28.2	12.4	34.7	21.8	27.8	19.8	190.6	164.7	89.8	74.4
MSE excluding the constant	10.6	5.8	19.3	13.3	29.3	15.0	9.8	4.5	42.2	16.8	26.6	20.8

* MSE : average value of (actual level - 95.0)^2, for the five β's.

V. CONCLUSION

We have suggested alternative pseudo-maximum likelihood estimators for Tobit models with first order serial correlation. These alternative estimators are not notably more costly to calculate than the pseudo-maximum likelihood estimators ignoring serial correlation, but they appear to be more efficient in terms of MSE.

We have also suggested a procedure for estimating the asymptotic covariance matrix of the APML estimators and setting confidence intervals. The true confidence levels associated with these intervals appear to be somewhat lower than the intended levels, but the MSE associated with these levels are not much higher than those associated with the maximum likelihood procedure that would be applicable if the latent variables were always observed.

Our approach could be extended to Tobit models with serial correlation of second or higher orders. For example, in the case of residual errors with a second order autoregressive scheme, the terms appearing in $\tilde{\mathcal{L}}^{**}$ would be :

a) the joint marginal density of the first Two Consecutive NonLimit (TCNL) observations;

b) the joint conditional density of all the nonlimit observations preceding the TCNL mentioned in a), given these TCNL observations;

c) the joint conditional density of each group of nonlimit observations preceded by the same TCNL observations, given these TCNL observations;

d) for every limit observation preceding the first TCNL observations, the integral of the conditional univariate density of this limit observation, given these first TCNL observations and the other nonlimit observations that precede these first TCNL observations;

e) for every other limit observation, the integral of the conditional univariate density of this limit observation, given the preceding TCNL observations and the following TCNL observations (if they exist) and given also all other nonlimit observations comprised between the above two sets of TCNL observations (or between the preceding TCNL and the end of the sample).

APPENDIX

Data used for the Monte Carlo Experiments

The variables are described in Section III, above.

observation number	C_{t-1}	C_t	M_{t-1}	B_{t-1}	K_t
1	1233.0	1364.0	99.0	230.0	1455.0
2	1364.0	1297.0	73.0	248.0	1507.0
3	1297.0	1359.0	98.0	340.0	1545.0
4	1359.0	1473.0	95.0	384.0	1637.0
5	1473.0	1436.0	117.0	353.0	1709.0
6	1436.0	1545.0	97.0	361.0	1731.0
7	1545.0	1646.0	126.0	331.0	1780.0
8	1646.0	1788.0	120.0	294.0	1887.0
9	1788.0	1867.0	152.0	297.0	2036.0
10	1867.0	1872.0	105.0	398.0	2163.0
11	1872.0	1899.0	161.0	342.0	2339.0
12	1899.0	1957.0	123.0	382.0	2436.0
13	1957.0	2064.0	284.0	333.0	2548.0
14	2064.0	2127.0	145.0	359.0	2582.0
15	2127.0	2201.0	204.0	305.0	2697.0
16	2201.0	2083.0	137.0	410.0	2762.0
17	2083.0	1960.0	220.0	350.0	2850.0
18	1960.0	1898.0	192.0	380.0	2924.0
19	1898.0	1785.0	237.0	332.0	2965.0
20	1785.0	1672.0	221.0	343.0	2947.0
21	1672.0	1578.0	269.0	308.0	2957.0
22	1578.0	1522.0	217.0	286.0	2923.0
23	1522.0	1686.0	223.0	232.0	2949.0
24	1686.0	1788.0	174.0	351.0	2932.0
25	1788.0	1962.0	247.0	399.0	2894.0
26	1962.0	1916.0	166.0	481.0	2860.0
27	1916.0	2096.0	281.0	357.0	2866.0
28	2096.0	2121.0	196.0	437.0	2837.0
29	2121.0	2363.0	278.0	347.0	2854.0
30	2363.0	2236.0	145.0	461.0	2814.0
31	2236.0	2259.0	280.0	526.0	2838.0
32	2259.0	2013.0	127.0	441.0	2828.0
33	2013.0	2042.0	287.0	538.0	2843.0
34	2042.0	1942.0	227.0	484.0	2807.0
35	1942.0	2110.0	292.0	428.0	2854.0
36	2110.0	2326.0	271.0	502.0	2889.0
37	2326.0	2403.0	269.0	499.0	2912.0
38	2403.0	2334.0	238.0	589.0	2887.0
39	2334.0	2445.0	272.0	536.0	2951.0
40	2445.0	2143.0	201.0	633.0	2903.0
41	2143.0	2150.0	270.0	718.0	2897.0
42	2150.0	2078.0	168.0	779.0	1949.0
43	2078.0	2177.0	143.0	677.0	2075.0
44	2177.0	1888.0	116.0	601.0	1959.0
45	1888.0	2063.0	131.0	557.0	2025.0
46	2063.0	2120.0	105.0	545.0	2076.0
47	2120.0	2294.0	164.0	441.0	2241.0
48	2294.0	2520.0	143.0	409.0	2715.0
49	2520.0	2692.0	199.0	403.0	2798.0
50	2692.0	2774.0	162.0	456.0	2787.0
51	2774.0	2804.0	166.0	465.0	2858.0
52	2804.0	2836.0	102.0	583.0	2861.0
53	2836.0	2886.0	140.0	569.0	2990.0
54	2886.0	3002.0	118.0	694.0	3064.0
55	3002.0	3080.0	177.0	697.0	3070.0
56	3080.0	3127.0	132.0	674.0	3094.0
57	3127.0	3227.0	168.0	587.0	3217.0
58	3227.0	3215.0	97.0	648.0	3301.0
59	3215.0	3274.0	143.0	688.0	3340.0
60	3274.0	3333.0	109.0	788.0	3413.0
61	3333.0	3346.0	151.0	826.0	3436.0
62	3346.0	3386.0	134.0	918.0	3416.0
63	3386.0	3521.0	183.0	844.0	3534.0
64	3521.0	3505.0	119.0	831.0	3532.0
65	3505.0	3737.0	160.0	799.0	3677.0
66	3737.0	3947.0	131.0	773.0	3804.0
67	3947.0	3951.0	138.0	623.0	3939.0
68	3951.0	3949.0	88.0	660.0	4044.0
69	3949.0	3951.0	139.0	705.0	4142.0
70	3951.0	4097.0	110.0	924.0	4282.0
71	4097.0	4021.0	211.0	1002.0	4467.0
72	4021.0	3853.0	152.0	986.0	4540.0
73	3853.0	3994.0	245.0	972.0	4851.0
74	3994.0	4115.0	201.0	943.0	4926.0
75	4115.0	4262.0	265.0	832.0	5009.0

REFERENCES

Beach, C.M. and J.G. MacKinnon, 1978, "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors", Econometrica 46, 1, 51-58.

Bussière, Luc, 1983, "Le modèle Tobit avec processus autorégressif d'ordre 1", M.Sc. thesis presented at Department of Economics, Université de Montréal.

Dagenais, M.G., 1964, "The Short-Run Determination of Output and Shipments in the North American Newsprint Paper Industry", Yale Economic Essays, fall 1964, 281-328.

Dagenais, M.G., 1982, "The Tobit Model with Serial Correlation", Economics Letters, 10, 3-4, 263-267.

Goldberger, A.S., A.L. Nagar and H.S. Odeh, 1961, "The Covariance Matrices of Reduced-form Coefficients and of Forecasts for a Structural Econometric Model", Econometrica 29, 556-573.

Gourieroux, C., A. Monfort et A. Trognon, 1982, "Estimation and Test in Probit Models with Serial Correlation", Working paper # 8220, CEPREMAP, Paris.

Lovell, M., 1961, "Manufacturers' Inventories, Sales Expectations, and the Acceleration Principle", Econometrica 29, 3, 292-314.

Mikhail, W.M., 1972, "Simulating the Small-Sample Properties of Econometric Estimators", Journal of the American Statistical Association 67, 339, 622-624.

Robinson, P.M., 1982, "On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables", Econometrica 50, 1, 27-41.

Robinson, P.M., A.K. Bera and C.M. Jarque, 1985, "Tests for Serial Dependence in Limited Dependent Variable Models", International Economic Review 26, 3, 629-638.

Tobin, J., 1958, "Estimation of Relationships for Limited Dependent Variables", Econometrica 26, 24-36.